

합성곱 신경망 빅데이터 학습을 통한 장소 이미지 식별 서비스 모델 개발*

장진욱 (농협대학교), 이동욱 (Jacobs University Bremen)

목 차

1. 서 론
2. 장소 이미지 수집처리
3. 합성곱 신경망 학습
4. 사물 인식 모델 학습
5. 웹 서비스 구축
6. 결 론

1. 서 론

최신 기술 중에는 컴퓨터에 빅데이터를 학습시켜 산업에 적용하는 사례를 여럿 볼 수 있다. 소비자들의 음성을 학습하는 인공 지능 비서 서비스 [1] 글의 맥락을 이해하여 번역하는 최신 번역기 기술 [2] 그리고 차로를 인식하여 주행하는 자율주행 모델 [3] 등을 예로 들 수 있겠다. 본 연구는 빅데이터를 이용하여 장소를 식별하는 이미지 학습 모델을 구축하고자 한다.

현재 사진 파일들은 사진을 촬영한 장소의 위치 정보(GPS)를 포함할 수가 있다. 하지만, 사진 촬영 기기에 따라 위치 정보를 저장할 수 없는 경우가 있으며, 저장이 가능하더라도 사용자의 고의로 위치 정보 저장 설정을 끄는 경우가 있다. 그 외의 경우에도, 위치 정보가 있는 사진을 여러 경로로

전송하거나 업로드 및 다운로드를 하는 과정에서 사진 속 위치 정보 데이터가 훼손될 수 있다.

현재와 같이 소셜 미디어가 발전한 사회에서는, 사람들이 서로의 사진을 보고 특정 사진이 촬영된 장소를 알고 싶어 하는 경우가 잦다. 하지만 위와 같이 사진에 위치 정보가 저장되어 있지 않다면, 사람들은 원하는 정보를 얻을 수가 없게 된다. 이 문제점을 시작점으로 본 연구는 합성곱 신경망 학습을 통해 사진이 찍힌 장소를 예측하는 테스트 모델을 구축하는 데에 목표를 두었으며, 온라인상의 사진 데이터들을 크롤링하여 학습함으로써 장소 예측의 정확도를 높였다.

2. 장소 이미지 수집처리

데이터 크롤링이란 특정 웹 페이지의 데이터를

* 본 논문은 2022년도 정부의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. NRF-2020R1G1A1005872)



(그림 1) 장소 이미지 수집, 학습, 예측 과정

수집하는 기법이다 [4]. 본 연구에서는 검색 포털들을 크롤링하여 합성곱 신경망 모델 학습에 사용할 장소 사진 데이터를 수집하였다. 장소로는 제주도의 5 명소인 오설록 티 뮤지엄, 만장굴, 몽상드 에일, 주상절리, 천지연 폭포를 선정하였다.

아래 (그림 1)은 크롤링한 사진들을 사용하여 모델을 학습 및 예측에 사용하는 전체 과정이다.

(그림 2)는 천지연 폭포의 사진들을 검색한 상태의 구글 모습이다. 본 연구는 검색 포털로서 네이버, 구글, 인스타그램을 사용하였으며, 각 검색 포털마다 크롤링 함수를 따로 작성하여 수집 성능을 개선하였다.

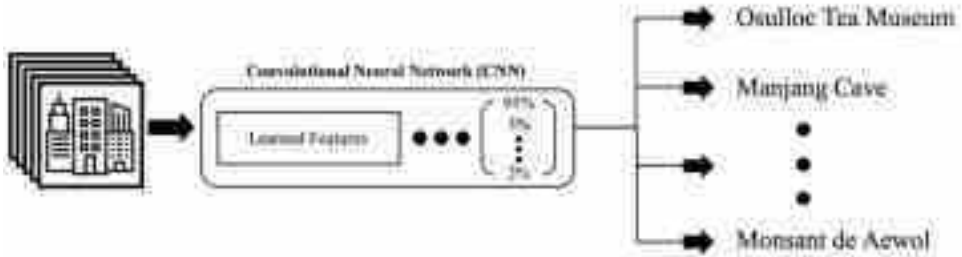
3. 합성곱 신경망 학습

합성곱 신경망은 사진 형식의 데이터를 학습하는 데에 최적화된 신경망 모델이다[5]. 특정 클래스에 속해있는 사진들의 공통 특징을 찾아 학습하고, 이후에 새로운 사진을 보았을 때 가장 비슷한 특징을 가지고 있는 클래스를 예측하여 결과로 출력하는 모델이다. 아래 (그림 3)은 합성곱 신경망을 이용한 학습과정 구조이다.

합성곱 신경망은 구글의 텐서플로 라이브러리를 활용하여 학습하였다[6,7]. 제주 5 명소에 대한 각 클래스 별로 대략 300개의 사진 데이터를 학습하였다. 그 결과로 학습 데이터 기준 예측 정확도 96.35%의 모델을 얻을 수 있었다. 아래 (그림 4)



(그림 2) 사진 검색 “구글 천지연 폭포” 결과



(그림 3) 합성곱 신경망 작동 구조

는 예측 정확도 테스트의 결과 창 캡처 화면이다. 각 클래스에 대하여 몇 개의 사진이 예측 성공하였는지 보여준다. 오설록 티 뮤지엄은 300개의 사진 중 292개, 만장굴은 300개 사진 중 297개, 몽상드 에월은 282개 사진 중 273개, 주상절리는 300개 사진 중 290개, 천지연 폭포는 300개 사진 중 274개가 예측에 성공하였다. 그 결과 합성곱 신경망을 학습함으로써 96.35%라는 예측 정확도의 모델을 도출하였다.

```
systemMemory: 8.98 GB
maxCacheSize: 2.67 GB

2022-03-08 00:56:39.183855: I tensorflow/core
2022-03-08 00:56:39.153150: I tensorflow/core
0, name: METAL, pci bus id: <undefined>
2022-03-08 00:56:39.274799: W tensorflow/core
2022-03-08 00:56:39.387221: I tensorflow/core
Class 0: 292 correct out of 300
Class 1: 297 correct out of 300
Class 2: 273 correct out of 282
Class 3: 290 correct out of 300
Class 4: 274 correct out of 300

***** Images test 2000 *****
Successfully Tested
```

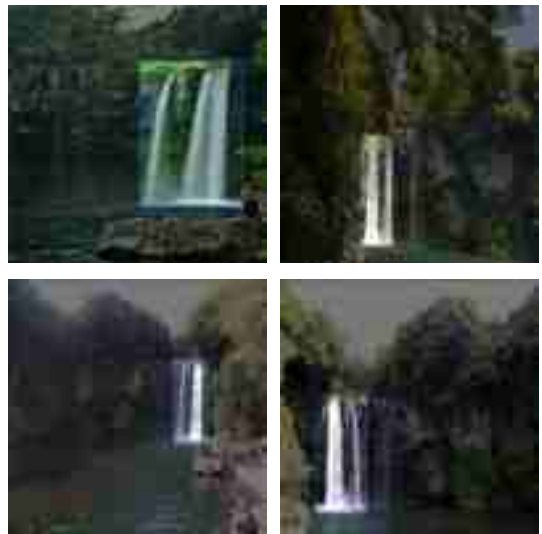
(그림 4) 합성곱 신경망 학습 후 예측 정확도 테스트 결과

4. 사물 인식 모델 학습

특정 클래스에 속한 사진들의 공통 특징을 찾아 분석하는 합성곱 신경망은 사물을 식별하는 데에 최적화되어 있는 모델이다. 이에 반해, 장소 사진

은 같은 곳임에도 불구하고 시간대에 따라 보여지는 모습이 달라질 수 있으며, 사람, 자동차, 동물 등 예측 불가능한 사물들이 함께 찍히는 경우가 있다. 또한, 촬영자의 위치와 각도에 따라서도 사진 속 특징이 달라질 수 있는 변수가 많다.

그에 따라 본 연구는 장소 사진 속에서 가장 공통 특징을 가질 확률이 높은 부분을 라벨링 하여 따로 학습시키는 방법을 모색하였으며, 사물 인식 모델을 학습시키는 방법을 선택하였다. 사물 인식 모델은, 사진 데이터를 학습하기 이전에 사진 속 특정 사물을 라벨링 하여 그 사물을 학습하는 모델이다[8]. (그림 5)는 사물 인식 모델 학습에 사용되는 사진 데이터의 예시이며, 본 연구는 제주 5



(그림 5) 천지연 폭포의 물줄기 라벨링

명소의 특징 부분을 각 클래스별로 1-2개 라벨링 하였다.

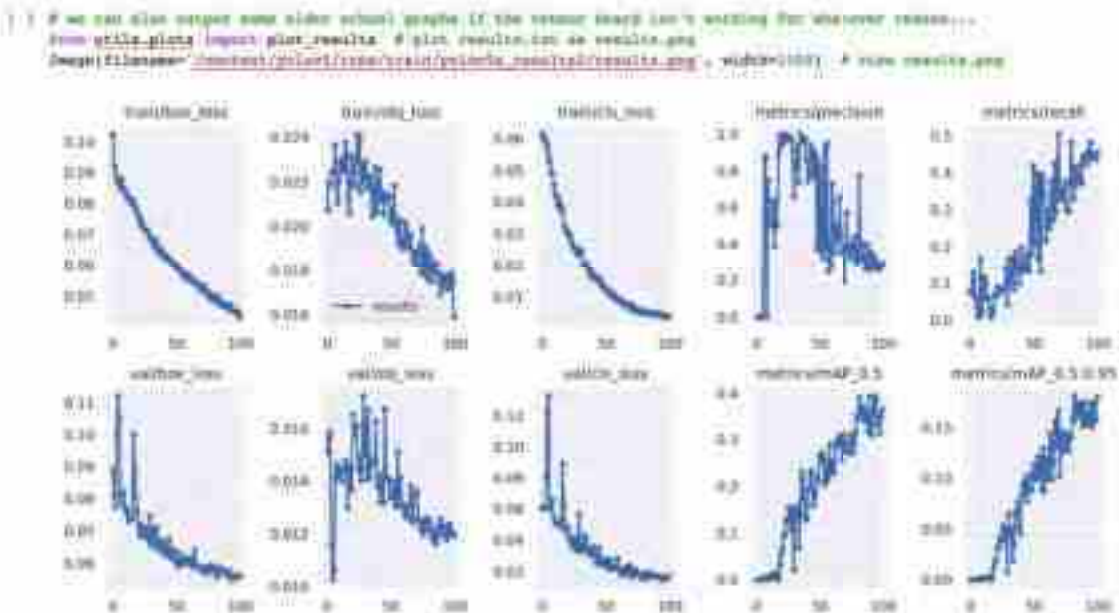
사물 인식 모델을 학습하는 데에는 클라우드 컴퓨팅을 사용하였다 [9]. 구글 코랩 (Google Colab) 은 구글에서 제공하는 무료 클라우드 서비스이다. 특히 서버를 통해 구글의 GPU와 TPU를 사용할 수 있기 때문에, 로컬에서 진행하기 버거운 연산들을 진행하는 데에 큰 도움을 준다. 사진 데이터는 무수한 픽셀값으로 이루어지기 때문에 모델을 학습하는 데에 많은 연산이 필요하며, 이에 따라 본 모델 학습은 구글 코랩을 사용하기에 걸맞은 상황이다. 아래 (그림 6)은 구글 코랩을 통해 학습된 사물 인식 모델의 분석 차트 화면 캡처이다. 각 그래프의 x축은 에포크(epoch), 즉 모델을 반복 학습하는 횟수를 의미한다. y축으로는 크게 손실률(loss), 정밀도(precision), 그리고 재현율(recall)이 분석되었다[10].

먼저 손실률이란 학습된 모델이 데이터에 얼마나 적합한지를 나타내는 수치로 낮을수록 학습 데

이터가 모델에 잘 적용된 것이다. 모델 분석 결과, 학습에 사용된 데이터들(train)과 모델에게 테스트할 새로운 데이터들 모두 에포크가 지날수록 손실률이 낮아져 최종적으로 “0”에 수렴하는 것을 볼 수 있다.

모델의 정밀도와 재현율은 아래 (그림 7)로 간단히 알아볼 수 있으며 모델 분석 결과 정밀도가 상승하였다가 에포크가 50이 지남에 따라 다시금 감소하여 35%에 머무르는 것과 재현율이 지속적으로 상승하여 45%에 머무르는 것을 확인할 수 있다.

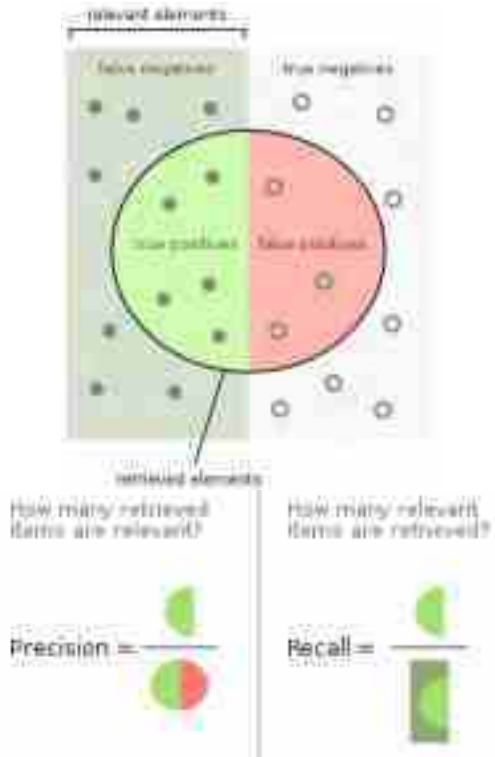
학습된 사물 인식 모델을 사용하여 사진을 분석하게 되면, 그 결과로는 바운딩 박스(Bounding Box)가 그려진 사진이 나오게 된다. 바운딩 박스란 사진 속에 검출된 특징 부분 혹은 사물이 있는 지를 찾아 그 위치를 그려주는 기법으로, 검출된 사물의 이름과 유사도를 출력한다. 아래 (그림 8)은 학습된 사물 인식 모델을 통해 몇 개의 장소 이미지를 분석해보았을 때 나온 결과 값들이며,



(그림 6) 장소 이미지 인식 모델 분석 결과

아래 결과를 통해서 우리는 입력된 사진들이 좌측

부터 순서대로 주상절리, 오설록 녹차 밭, 천지연 폭포, 만장굴이라는 것을 알 수 있다.



(그림 7) 정밀도와 재현율 시각화

5. 웹 서비스 구축

본 연구는 웹 서비스를 개발하여 사람들로 하여금 쉽게 장소 이미지 식별 모델에 접근할 수 있도록 하였다. 웹 개발은 장고(Django) 라이브러리를 사용하여 개발되었다. 아래 (그림 9)은 웹 사이트의 메인 페이지이다.

메인 페이지의 중앙에 위치해 있는 검색 바는 이미지 파일을 업로드 할 수 있도록 제작하였다. 이미지를 검색 바 위로 드래그하거나 옆에 있는 사진 버튼을 눌러 파일 업로드를 하면 바로 백엔드에 연결되어 있는 합성곱 신경망 학습 모델에 사진 파일이 넘어간다. 사용자가 사진을 입력하면 웹 페이지는 합성곱 신경망 모델을 사용해 장소 사진을 예측하여 그 결과를 클래스 번호로 출력한다. 예를 들어 5개의 장소 클래스를 학습한 본 서비스의 모델은 예측 결과로 0-4중 하나의 클래스



(그림 8) 입력된 장소 사진 (상), 사물 인식 모델 예측 결과 (하)



(그림 9) 장소 이미지 식별 사용자 화면

번호를 출력한다. 본 연구는 0-4번의 클래스 번호를 오설록 티 뮤지엄, 만장굴, 몽상드 애월, 주상절리, 천지연 폭포 순으로 지정하였다. 아래 (그림 10)은 사진을 업로드 하는 두 가지 방법이다.

그 이후, 모델에서 클래스 번호를 예측값으로 받은 웹 서비스는 MySQL에 연결된 장소 클래스 관련 정보를 받아온다. MySQL에는 5개의 장소

이미지 클래스에 대한 테이블이 생성되어 있으며, 각 클래스에 대하여 장소 이름, 장소 대표 사진, 장소 설명 글이 정보로 저장되어 있다. 아래 (그림 11)은 MySQL에 저장되어 있는 제주 5 명소의 정보이다.

사용자의 사진을 입력받고 합성곱 신경망을 통해 예측을 한 후 결과로 나온 클래스 번호를 MySQL의 장소 클래스 인스턴스들과 매칭시켜 정보를 받아온다. 그리고 메인 웹 페이지는 사용자에게 검색한 이미지 장소에 관련된 정보를 출력한다. 이 과정을 비추어 보아 본 모델은 추후 사용자가 직접 자신이 선호하는 장소 이미지 클래스를 학습시키는 사용자 맞춤 장소 식별 모델로써 쓰일 수 있다.

아래 (그림 12)은 검색에 사용된 이미지와 검색 결과로 나온 메인 웹 페이지의 모습이다. 추가적으



(그림 10) 사진 드래그 (좌), 사진 아이콘 클릭 (우)



(그림 11) MySQL에 저장된 5개의 장소 클래스 리스트



(그림 12) 검색 이미지 (좌), 웹 서비스 모델 예측 결과 출력 (우)

로, 사용자가 입력한 이미지는 추후 전처리 과정을 거쳐 합성곱 신경망 모델을 성장시키는 데에 데이터로 쓰기 위해 서버 파일 디렉토리에 저장된다.

6. 결 론

본 연구는 위치 정보가 저장되어 있지 않은 사진 데이터들에 대하여 픽셀값을 분석함으로써 사진이 촬영된 장소가 어디인지 예측하는 서비스를 구축하였다. 그 결과 합성곱 신경망을 학습함으로써 96.35%라는 예측 정확도의 모델을 도출하였으며 현재 웹 서비스 또한 이 모델을 기반으로 돌아가고 있다. 추가적으로 사물 인식 모델을 학습하여 장소 이미지 식별을 테스트하였으나, 아직까지는 합성곱 신경망의 예측 정확도로 사용자들에게 필요한 정보를 줄 수 있을 것이라 예상된다.

추후 장소 클래스 개수가 늘어나고 학습할 데이터의 양이 늘어남에 따라 사물 인식 모델을 웹 서비스에 통합할 수 있다. 시스템 개발자 관점에서 장소 클래스를 추가함에 따라 데이터 학습량이 늘어날 것이며, 서버상에 저장된 사용자들의 입력 사진들이 데이터로 사용되는 부분에서 학습량이 더욱 늘어날 것이다. 그리고 데이터가 늘어남에 따라 장소 이미지의 다양성이 드러날 것이기 때문에 사물 인식 모델의 사용이 필요해질 것으로 예

상한다.

결과적으로 합성곱 신경망 모델과 사물 인식 모델을 효율적으로 함께 사용하기 위한 결정 트리(decision tree)를 고안한다면 사용자들에게 더 많고 정확한 정보를 주는 웹 서비스로 발전될 것이다 [11]. 본 서비스의 장소 이미지 딥러닝 모델은 입력된 장소를 기준으로 주변 관광지, 유명명소, 음식점, 숙박 장소 등 장소 정보를 제공하는 다양한 서비스로 확대 적용 가능하다.

참 고 문 헌

- [1] 김동주, 이가람, 김호원, 시를 활용한 음성인식 기법 및 개인비서 서비스 기술 분석, 한국통신학회 학술대회논문집, 213-214, 2017.
- [2] 김재원, 번역에서의 맥락과 맥락효과에 대한 연구, 미래영어영문학회 학술대회 자료집, 171-178, 2018.
- [3] 김영민, 김형수, 자율주행 차량의 도로 평면 선형 기반 차로이탈 허용 범위 산정, 한국ITS학회 논문지, 15(4), 81-90, 2016.
- [4] Cem Dilmegani, "What is Web Crawling? How it works & Examples", Available online at: <https://research.aimultiple.com/web-crawler/>, 2022.
- [5] Prabhu, "Understanding of Convolutional

Neural Network (CNN) Deep Learning”, Available online at: <https://medium.com/@RaghavPrabhu/understanding-of-convolutional-neural-network-cnn-deep-learning-99760835f148>, 2018.

- [6] Chatterjee M, “What is TensorFlow? The Machine Learning Library Explained”, Available online at: <https://www.google.com/amp/s/www.mygreatlearning.com/blog/what-is-tensorflow-machine-learning-library-explained/%3Famp>, 2020.
- [7] Tensorflow, “Customization basics: Tensors and operations”, Available online at: <https://www.tensorflow.org/tutorials/customization/basics>, 2020.
- [8] Lihi Gur Arie, “The practical guide for Object Detection with YOLOv5 algorithm”, Available online at: <https://towardsdatascience.com/the-practical-guide-for-object-detection-with-yolov5-algorithm-74c04aac4843>, 2022.
- [9] Bryan White, “Google Colaboratory: Python in the Cloud”, Available online at: <https://medium.com/swlh/google-colaboratory-python-in-the-cloud-1cefd1cbe7e>, 2020.
- [10] Jason Brownlee, “How to Calculate Precision, Recall, and F-Measure for Imbalanced Classification”, Available online at: <https://machinelearningmastery.com/precision-recall-and-f-measure-for-imbalanced-classification/>, 2020.
- [11] Anshul Saini, “Decision Tree Algorithm – A Complete Guide”, Available online at: <https://www.analyticsvidhya.com/blog/2021/08/decision-tree-algorithm/>, 2021.

저 자 약 력



장 진 욱

이메일 : jjw@nonghyup.ac.kr

- 2013년 건국대학교 신산업융합학과 (박사)
- 2001년~2005년 국방부 정보사령부 전산장교
- 2011년~2013년 SK communications CTO PMO
- 2021년~현재 농협대학교 협동조합디지털경영과 교수
- 관심분야: 장소 이미지 딥러닝, 디지털전환, 소프트웨어 품질



이 동 욱

이메일 : dongwooklee1201@gmail.com

- 2022년 Jacobs University Bremen Intelligent Mobile Systems (학사)
- 관심분야: 컴퓨터 비전, 머신 러닝, 로봇 공학